LLNL-TR-649415

LAWRENCE LIVERMORE NATIONAL LABORATORY

A Probability-based Approach to Evaluating Security of Nuclear Fuel Cycles

A. D. Lamont, T. A. Edmunds

February 6, 2014

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Executive Summary

The National Research Council has reviewed several studies of the security and proliferation resistance of nuclear fuel cycles. The study found several recurring weaknesses in the execution of a commonly used method, which they call *predefined frameworks*. The predefined framework studies reviewed were attempts to utilize multiattribute utility theory to score and rank alternative fuel cycles. This white paper indicates how some of these weaknesses in the pre-defined frameworks (MAU) evaluation process can be addressed using more rigorous MAU assessment methods and models. First, the paper suggests how the weaknesses in previous MAU-based assessments could be corrected in future studies. The remaining sections of this paper describe an alternative approach that is based upon both probability theory and a multiplicative form of a multiattribute utility function.

Table of Contents

1.0 Introduction

The National Research Council (NRC) has reviewed two methods for evaluating the security and proliferation resistance of nuclear fuel cycles [NRC 2013]. The first method is referred to as *case-by-case assessments*. Under this approach, a panel of subject matter experts is convened to examine alternative fuel cycles and render an overall opinion as to their security or proliferation resistance. The second approach is to use "pre-defined frameworks" for the evaluation. These frameworks identify the features of the process to be evaluated, provide scales to evaluate the features, and usually provide a method to aggregate the individual evaluations into an overall figure of merit (FOM) for the process. They found several recurring weaknesses in these predefined frameworks, which are identified below.

- In some cases there is a confusion between proliferation resistance and proliferation risk. Proliferation resistance measures the likelihood that an adversary will fail, given that an attempt is made. Proliferation risk measures the probability that an attempt will be made <u>and</u> that it will succeed.
- The figure of merit produced by these frameworks does not have an unambiguous interpretation that is relevant to policy decisions. Some approaches claim to be based on multi-attribute utility (MAU) theory. The correctness of the MAU implementation for several of these examples is highly questionable. However, even if done correctly the FOM measures the <u>relative</u> desirability of a given configuration. But, this does not state how the desirability is measured. It is simply a preference between configurations. Other approaches use a fairly arbitrary weighting and rating scheme that cannot be interpreted clearly in terms of an objective measure. If a method is not based on some measure that is external to the evaluator, then it is difficult to draw any conclusions about whether or not improvements should be made, or that the process is 'over protected' and less should be spent on protecting it.
- The methods and procedures used to calibrate the FOM are not based on clear principles (i.e. theories). They often rely on expert opinion that is elicited in ad hoc methods.
- The methods do not address the uncertainty associated with the assessments partly due to the fact that they are not defined in terms of probability and partly due to the fact that the overall FOM is based directly on quantitative descriptions of the process. There is no step that assesses the implications that the process description has for the probability of proliferation.
- The structure of the analysis does not lend itself to a good understanding of why the overall scores came out the way they did and does not lend itself to diagnosis of specific problems. These methods do not allow us to identify the most cost effective improvements in a process.

This white paper indicates how some of these weaknesses in the pre-defined frameworks (MAU) evaluation process can be addressed using more rigorous MAU assessment methods and models. The next section suggest how the weaknesses in previous MAU-based assessments could be corrected in future studies. The remaining sections of this report describe an alternative approach that is based upon both probability theory and a multiplicative form of a multiattribute utility function.

2.0 Solutions for NRC criticisms of previous pre-defined frameworks

The NRC's review of previous studies identified several shortcomings. As indicated in this section, these shortcomings could be overcome by taking relatively simple measures.

2.1 MAU execution and theoretical foundations

The NRC report identifies shortcoming in *execution* of the pre-defined framework (MAU) processes that were reviewed [NRC 2013, pg. 8], but it does not state that the *theoretical foundation* of MAU is unsound. In addition, they state that the use of expert opinion and knowledge is clear and understood by policymakers while use of predefined frameworks is not clear (a black box) [NRC 2013, pg. 2]. Finally, they found that the frameworks studied were not used to inform policy decisions. It is not surprising that processes that were poorly executed were not understood by policymakers and were not effective in informing policy decisions.

2.2 Consistency of MAU-based assessments

The NRC study recommends use of expert panels convened to support evaluations of fuel cycles (case-by-case assessments) rather than a MAU approach. While case-by-case assessments do offer a large measure of freedom for experts to explore and rank fuel cycles, results of the process may not be repeatable and the logic underlying the ranking may not be rigorously traceable. In contrast, the MAU approach captures, quantifies, and archives expert options in the form of utility functions and weights to combine them. When properly executed, the reasons behind the rankings can be traced to assumptions and subjective value judgments provided by subject matter experts and decision makers. Moreover, the sensitivity of fuel cycle rankings to such judgments and assumptions can be explored using the MAU model. The model becomes a platform for dialog among experts and stakeholders. It can identify which issues are important (affect the ranking) and which are not so that debate can focus on only the crucial issues. Finally, the MAU provides a tool for fuel cycle designers. New fuel cycle designs can be developed and scored using the MAU function to compare to exiting cycles. Sensitivity analysis provides a mechanism to guide developers towards more effective fuel cycle solutions.

2.3 Expert elicitation process

The NRC study identified three shortcomings in the elicitation processes used in the studies review. First, the processes were deemed faulty due to overreliance on surveys to collect data [NRC 2013, pg. 1]. Second, the elicitation processes were not well documented [NRC 2013, pg. 10]. Third, the expert elicitation sessions did not include a bona fide decision analyst trained to conduct such sessions and interpret results [NRC 2013, pg. 35]. All three of these shortcomings could be overcome by using established and documented procedures implemented by experienced decision analysts.

2.4 Uncertainty and sensitivity analysis

It is a recognized principle in decision analysis that a decision support model should not generate a single, final answer and not replace the decision maker. The primary purpose of building a

MAU model or other decision support tools is to exercise the model with many different parameter values in order to assess the impacts of uncertainties and to explore the sensitivity of decisions to subjective judgment and assumptions. The key outputs of the model are qualitative insights regarding policies and investment decisions. The NRC study noted shortcomings in previous studies in this regard. [NRC 2013, pp. 1, 10, and 36]. A wide range of methods and decision support software tools are available to support uncertainty and sensitivity analyses.

2.5 Adversary modeling

The NRC study indicates that the shortage of data relevant for proliferation and theft threats to nuclear fuel cycles is a significant challenge to identification and quantification of threats to nuclear fuel cycles [NRC 2013 pg. 10]. Although this does pose a challenge to validating models of threats and risks, bounding cases can be developed by assuming an optimizing adversary with perfect knowledge about the fuel cycle and its safeguards and security procedures. Decision analysis techniques such as decision trees can be used to identify weak points in the fuel cycle. Monte Carlo simulation of the decision trees can then be used to model adversaries with incomplete information about the system they intend to attack or exploit. Some methods are discussed in [Maurer 2009, Ni 2013].

2.6 Application to physical security

The NRC study indicates that physical security against theft or diversion of weapons usable nuclear material from fuel cycles is not within the scope of their study [NRC 2013, pg. 16]. However, we assert that the same MAU methods can also be used to assess the vulnerability of nuclear fuel cycles and processes to theft. Past studies of nuclear fuel cycles have used MAU methods to assess both proliferation resistance and theft [Ward 2007].

3.0 Desirable Characteristics of an Evaluation Framework

The evaluation framework provides a numerical measure of the relative desirability of different configurations of the fuel cycle and a process. Various methods have been proposed for constructing such a measure. Each of the methods scores the features of the process. In the terms used here, a “score” is an objective description of one aspect of the process (e.g. the mass of material in the process, or the specific properties of the material). Based on the objective scores, a method is applied to convert the scores into an overall “figure of merit”. Presumably, the figures of merit for two different processes, or two different versions of a given process, can be compared to determine which one is superior.

3.1 Analysis results should use measures meaningful for making decisions

Evaluation methods are implemented in order to guide decisions. They should tell us if a fuel process is not safe enough and money should be invested to make it safer by adding safeguards or by moving to a different process that may be more costly but is safer. An evaluation framework produces a numerical value (or a set of values) that characterized the desirability of the process. If these are to be useful for decision making one of two conditions must hold: Either the numerical results include all of the issues that are relevant for making a decision, or, the numeric value must be in terms of measures that can be directly compared to the other relevant measures that are not included in the evaluation. We can provide two examples.

First, some analytical frameworks mentioned in the NAS report use multi-attribute utility. Normally a utility function is used to measure the relative desirability of different outcomes, as viewed by a decision maker. In that case, the utility function models the decision maker's preferences in terms of his willingness to trade-off one aspect of the problem to improve another aspect. When correctly implemented a multi-attribute utility function includes all relevant aspect of the problem. The examples of utility function based frameworks cited by the NAS do not include costs. The resulting utility value cannot be directly compared to costs since the utility measure is just a measure of relatively preferences among the issues that are included in the function. The utility function may tell us that A is better than B. But is does not tell us how much better. Most important it does not tell us if A or B is good enough, or if it would be preferable to spend more money to improve from B to A.

Second, the evaluation function can be defined in terms of a commonly understood, objective measure. In the discussions below it is suggested that the evaluation function be defined in terms of the probability that an adversary would fail to obtain material and fabricate a device, given an attempt. With that definition, a decision maker can compare alternative processes and measure the difference between them. And a decision maker can compare the benefits of improving a process to the costs of such improvements.

3.2 Analysis should be diagnostic

It should be possible to identify specific feature so the fuel cycle that determine the overall assessment. If the overall measure and sub-measures are defined in meaningful terms, it should be possible to determine in an objective way the relative importance, or value, of improving different parts of the system.

The value of a method as a diagnostic tool is increased if the method allows us to divide up the problem into physically meaningful components, model the components, and then combine them into an overall measure. The process of simply defining and modeling the components often provides clearer insights to the problem and a better understanding of which factors determine the results. By dividing the analysis into components we can better understand which components influence the results and which have relatively little effect on our overall evaluation.

3.3 Judgments used should be well defined and transparent

We ask subject matter experts (SMEs) to make judgments about impact that features of the process have on the overall evaluation. These judgments should be defined such that it is objectively clear what is being assessed. That is, two SMEs might not agree on the assessed value, but they do agree about the nature of the feature or event they are asked to assess.

In this case we will be asking SMEs to assess probabilities of adversary failure at different stages of a process and under different conditions. The framework should make it clear for the SME what the nature of the adversary is, what the adversary must accomplish for success, and what obstacles might prevent success.

4.0 An Alternative Framework Based in Probability Theory

4.1 Key elements

- A metric that is based on the probability of adversary failure, given attempt. This probability measure is a basic building block of any analysis.
- The analysis is built up in levels and explicitly evaluates component of the process. This allows us to identify the contributions from each component and diagnose the strong and weak points of a process.
- We need to use scales that objectively describe features of the fuel cycle to estimate probabilities of key events in proliferation by an adversary.
- At a certain level in the analysis there may be several features that bear on the probability that an adversary will fail in accomplishing an element of his strategy. However, the exact way that they combine to affect the overall probability cannot be easily modeled using standard probability models. In these cases we need a reasonable model that can represent the SME's assessment of probability of adversary failure, given the implementation of the features.
- When using subjective assessments several conditions must hold:
    - The events being assessed must be clearly defined in a way that different people will agree about the nature of the event being assessed. They might not agree on the probability, but they should have a common understanding of the definition of the event
    - The method of combining the effects of the various factors must at least be amenable to clear, defensible probability assessments in a set of bounding cases. The model allows for interpolation between these bounding cases. We require that the interpolation be reasonable.
    - Since there is a clear probability interpretation at each level with a clear definition of the events to be assessed, the subjective probability model can be replaced with more standard probability models, if there are resources available to conduct the modeling.

4.2 Fuel cycle performance characterization

From the point of view of protection against proliferation, we clearly wish to reduce the probability that material is obtained and fabricated into a device and to mitigate the consequences if a device is set off. We can divide the question into components:

- Successfully obtaining material and converting the oxide into a metal.
- Successfully fabricating a device
- Setting off the device, and
- Consequences

For simplicity, this discussion examines the analysis of the first component. The approach can be extended to include the other components.

This discussion focuses on the event that an adversary is able to obtain material from a processing facility, and fabricate a device. In principle, the adversary could be an insider or an outsider. In either case, the adversary must enter the facility, acquire the material, remove it from the facility and take it to a location where he can work on it, process it into suitable materials, and fabricate a device. The overall figure of merit proposed here is a measure of the probability that an adversary would fail to acquire material and produce a working device.

In this discussion it can be assumed that the adversary has full information about the facility (this is the design basis threat) and that he attacks the weakest points in each stage of the strategy. Conceptually, it is desirable to examine all possible strategies to find the best one from the adversary's point of view. This could be done, in principle, using an approach such as the ASSESS model [Al-Ayat 1993]. In many cases that will be impractical due to either time and cost limitations or due to the fact that facility is not yet fully designed. In addition to modeling adversaries with perfect information, one could also define other adversary characteristics and strategies. These should be made clear at the start of the analysis process.

This figure of merit does not address the probability that a particular adversary would attempt a theft at a given facility or a given material. It only addresses the probability of adversary failure, given that the attempt is made. In this form, the figure of merit is a building block for more extensive analyses. If the probability of failure is considered to be an input to the adversary's decision process in either deciding to make an attempt or choosing a target, then this calculation is a basic step in the higher-level analysis.

There generally will be multiple different strategies that and adversary could use at each step: he might surreptitiously acquire material, or he might falsify documentation that allows him to openly acquire and remove it. The probability of success at each step depends on the strategy chosen by the adversary.

The model proposed below is fundamentally structured to assess the probability of adversary failure for a single adversary strategy. This provides a well-defined set of questions that an SME can answer regarding the probabilities of adversary failure. The overall analysis effort, however, is concerned with the overall probability of adversary failure considering all of the strategies that an adversary might attempt. To evaluate this overall probability correctly, the analyst should assess the probabilities of adversary failure for each strategy and assess the probability that the adversary would choose each strategy. Then the overall probability of adversary failure can be computed as the sum of the probabilities of choice times the probabilities of failure.

5.0 Multiplicative Multiattribute Framework Description

An analysis of proliferation resistance will have several levels. The lower levels assess the probabilities that the adversary can successfully evade sets of safeguards or accomplish tasks. The upper levels roll up all of the probabilities over the entire strategy to arrive at an overall FOM.

We have observed that the highest level of the framework can be constructed as a well-defined probability calculation where several events clearly must happen. Given the probabilities of those events, we can use conventional probability methods to compute the overall figure of merit as a probability.

At lower levels in the framework, we find that there may be several features of the process or safeguards that clearly affect the probability of adversary failure in accomplishing an essential task, but the exact way that these features affect the probability is not clear without a careful modeling of the facility and the adversary strategy. The assessment of the effect of a single feature may be comparatively easy to address. The conceptual difficulty lies in determining how the features might interact to influence the probability of adversary failure. We have taken the approach of applying a functional form that can plausibly represent the combinations of features. We can assess probabilities from an SME for cases that are relatively simple and clear. These responses are then used to calibrate the overall function. Again, there are methods to compute these probabilities more rigorously. But they require extensive modeling of the facility and the adversary strategy.

There is a strong parallel between the form of the equations for multi-attribute utility modeling and a probabilistic risk analysis. We can exploit some of the insights from utility theory to help build the model.

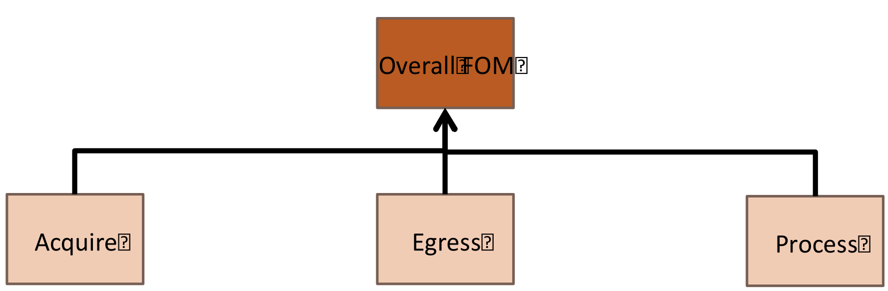
A complete analysis might include tasks such as entrance to the facility, fabrication of a device, and setting off the device in a location that will cause great harm. For simplicity, we will just focus on these three steps:

- Acquire the material: actually get his hands on it and remove it from where ever it is stored
- Egress from the facility and transport the material to some secure working location. We assume that if he is able to successfully egress from the facility, he can then move the material to a secure location
- Process the material to make it suitable for a device. This depends on the composition of the material that is stolen. Generally would be a process of chemical separation and purification.

5.1 Objectives hierarchy

At the top level the adversary must accomplish the distinct stages shown in **Figure 1** in order to successfully create a device.

Figure 1: Top level stages that the adversary must complete



If we have the probabilities that the adversary will fail to complete each of these stages, we can calculate the probability of overall adversary failure as

$$\textit{Probability of adversary failure} = 1 - (1 - P_{acq}^{fail}) \bullet (1 - P_{egr}^{fail}) \bullet (1 - P_{prc}^{fail}) \quad (1)$$

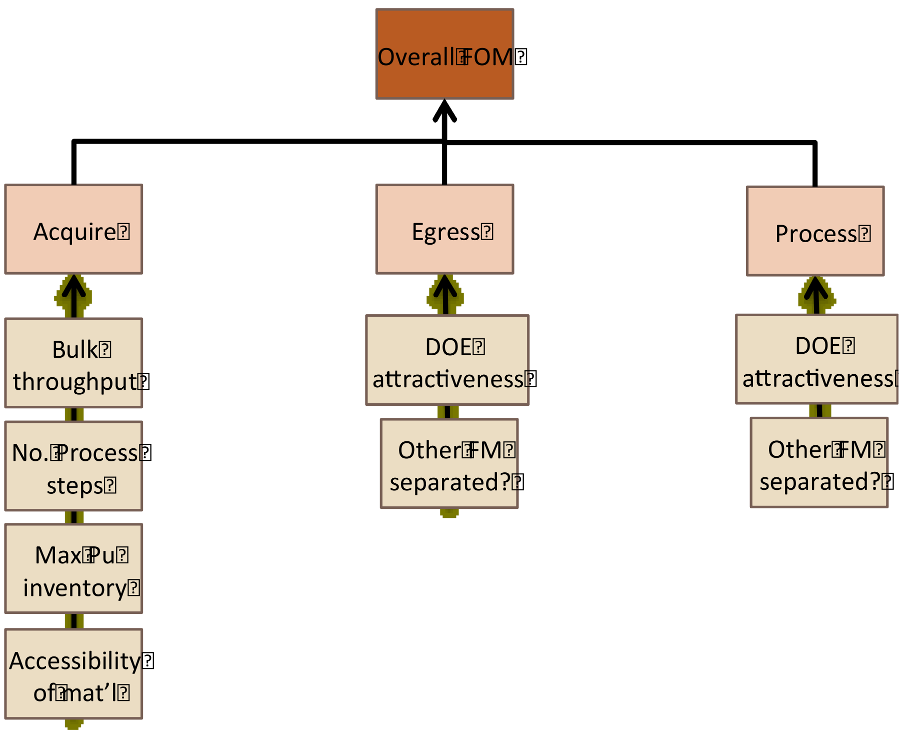
Where:

P_i^{fail} = probability that the adversary fails at stage i

5.2 Full structure organized to facilitate expert elicitation of probabilities

The probability of adversary failure at each stage is a function of the safeguards and properties of the fuel cycle. The first step in developing elicitations is structuring the remainder of the problem in a form that lends itself to probabilistic elicitation. **Figure 2** illustrates the full analysis.

Figure 2: Structure of the full analysis



Under each of the major stages (acquire, egress, and process) the features of the relevant safeguards and fuel cycle characteristics are listed. This example uses the safeguards and process features defined by Dyer et al [Dyer 1997]. The features are:

- Bulk throughput
- Number of process steps,
- Maximum Plutonium inventory,
- Accessibility of material,
- Type of nuclear accounting system,
- DOE attractiveness level,

- Whether or not other fissile materials are separated in the process.

Dyer et al [Dyer 1997] have defined scales for each of these features to describe the level of the implementation or the intrinsic characteristics of each feature at a given site and fuel cycle. These will be used in this example.

We note that the same feature can appear in several places in the analysis. This is due to the fact that a given feature might be relevant to several processes within the analysis. In the analysis, the implementation of the feature is constant across the entire analysis (it is only implemented once in the facility), but the impact of a given level of implementation of the feature might be different in different part of the analysis.

An adversary might be stopped at a step due to the fact that it is simply too difficult to complete (e.g. the material is well secured in an inaccessible storage), or because the adversary's actions are detected and the adversary is stopped. In structuring an analysis it may be helpful to separate these two reasons for adversary failure. **Figure 3** presents a possible extension to the analysis that separates these issues. This approach has structured the analysis in each major stage in terms of "Difficulty" and "Detection". Features that make detection more likely are the accounting system, simple observation by other workers, the nature of the material may have a detectable signature, or there may also be other fissile materials available that an adversary could use to mask the signature of the material being stolen. "Difficulty" refers to the physical barriers to actually accomplishing the task. Some features of the process or safeguards may simply make it physically difficult to accomplish, even without detection and prevention by the facility staff.

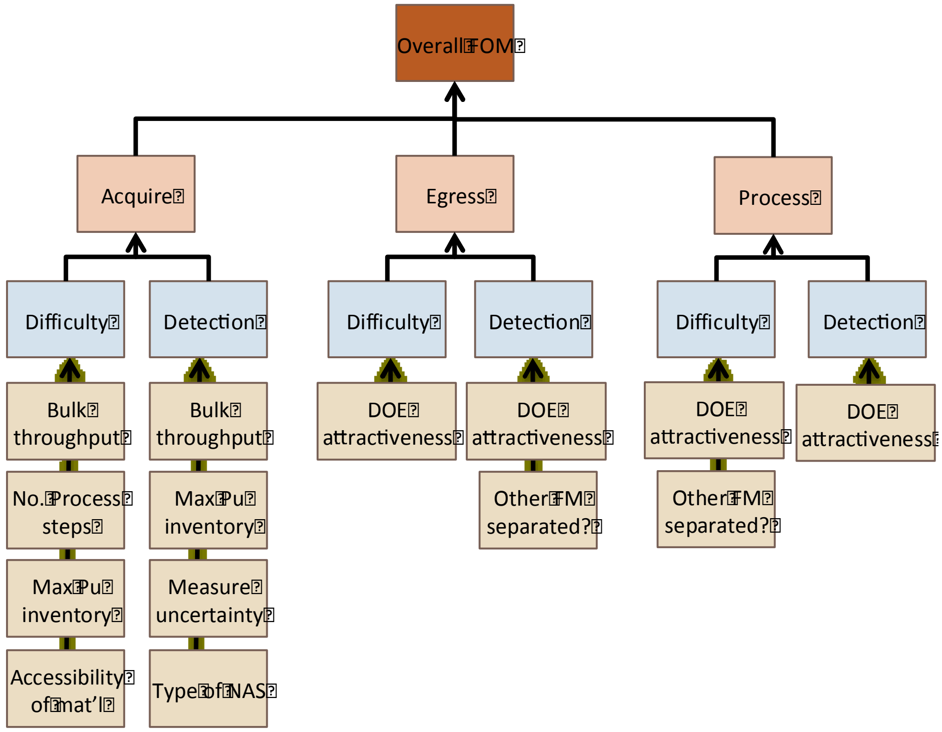
It is not logically necessary to use this structure of "Difficulty" and "Detection". It is proposed here as an aid to organizing the analysis and the assessments, but the assessments could be structured without using these features.

The rest of this discussion will describe the proposed method for estimating the probability that an adversary will fail in a stage due to either Difficulty or Detection. In each area we will first assess the probability that the adversary fails because the theft is simply difficult, and then assess the probability that he fails because he is detected and prevented. To succeed, the adversary must succeed at both phases of the stage. Then a calculation similar to equation (1) can be used to compute the probability of failing at the stage.

The model is based on probability calculations using expert elicitation. It should be emphasized that at each part of the analysis where elicited probabilities are used, more elaborate modeling or simulation can, at least in principle, provide probabilities that are explicitly based in probability theory.

After the discussion of eliciting low level probabilities, the method of rolling up the entire result using equations of the form of eq (1) will be described. As is noted earlier, there are close analogies between multi-attribute utility theory and the probability calculations described here. The insights that can be gained from utility theory are also discussed.

Figure 3: A possible extension of the model to separate the difficulty of accomplishing a step from the possibility that the adversary will be detected and stopped



5.3 Modeling probability of adversary failure

To illustrate the process, we will discuss the steps needed to compute the probability of adversary failure to acquire material due to difficulty. The analysis requires the following modeling steps:

- Describe the features of the process and facility
- Convert these descriptions into a numerical value that is related to the probability of adversary failure at each step (there can be different conversions for different steps.)
- Define a functional form that can combine the scores for all of the features that bear on a step
- Calibrate the functional form through expert elicitation

Once the function for "difficulty" is calibrated, the functional form for "detection" is calibrated. These give the probabilities of adversary failure for both phase of the "acquire" stage. These can be combined to determine the probability that the adversary would fail in the "acquire" stage.

5.3.1 Describe features of process and facility

For this illustration the features of the process that bear on that probability for this example are:

- Bulk throughput
- Number of process steps
- Maximum Plutonium inventory

- Accessibility of material

Each of the features will be characterized for the process and facility. These are measured according to observable, objective features of the process.

- Bulk throughput: measured in metric tons per year of bulk material processed. Note that a facility may have zero tons of bulk material, but substantial amounts of material processed as discrete items.
- Number of process steps: A processing step was defined as an action or activity that involves a form change of greater than one percent in the physical or chemical properties of the material.
- Maximum Plutonium inventory: Ranges from 0 to 50 MT
- Accessibility of material: The measure is really a combination of three factors: the accessibility of the plutonium in process, the accessibility of the "container", and whether special handling equipment is required to move the plutonium. **Table 1** provides a description of each level of characteristics.

Table 1: Measures to describe Accessibility of Material [Dyer 1997, Table 5]

Table 5 - Combinations for Accessibility of Plutonium "Sub-measures"

Accessibility of Pu		*Accessibility of Container*		*Special Equipment to Move Pu*	
touchable Pu (T) tamper-indicating container (C)		hands on container (H) remote / robotics (R)		yes (Y) no (N)	
Measure	Score	Measure	Score	Measure	Score
C	1	R	1	Y	1
C	1	R	1	N	0
C	1	H	0	Y	1
C	1	H	0	N	0
T	0	R	1	Y	1
T	0	R	1	N	0
T	0	H	0	Y	1
T	0	H	0	N	0

5.3.2 Convert descriptions into values related to probability of adversary failure

Each of these features bears on the probability that an adversary would fail to accomplish a task. The relationship between the value of the measure and the probability of adversary failure is not necessarily linear. Often there is some sort of threshold effect. For example, if the measure is below some threshold value there is little effect, and a strong effect when it exceeds that value. A value function is defined over each of these features to relate the score value to a value function that is linearly related to the probability of adversary failure.

The value function is scaled to range of 0 to 1 (the figures below are scaled from 0 to 100%). This is not interpreted as the probability itself. It is only required to be linearly related to the probability. To say that it is linearly related implies that if the value function doubles in value, the probability doubles in value. Note that the sections below point out that the overall functions are scaled so that the probability of adversary failure does not exceed 1.0. Example value functions for Bulk Throughput, Number of Process Steps, Maximum Plutonium Inventory, and Accessibility of Material are shown in **Figure 4** through **Figure 7**. These are taken from Dyer et al.

Figure 4: Value function for Bulk Throughput

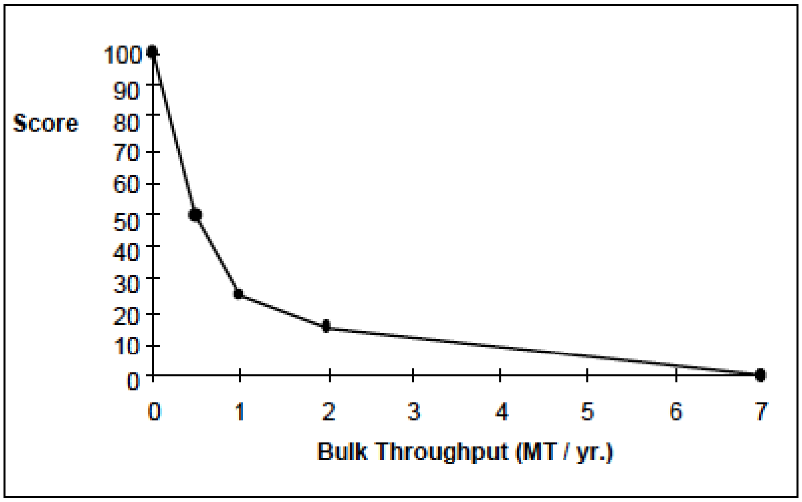


Figure 5: Value function for Number of Process Steps

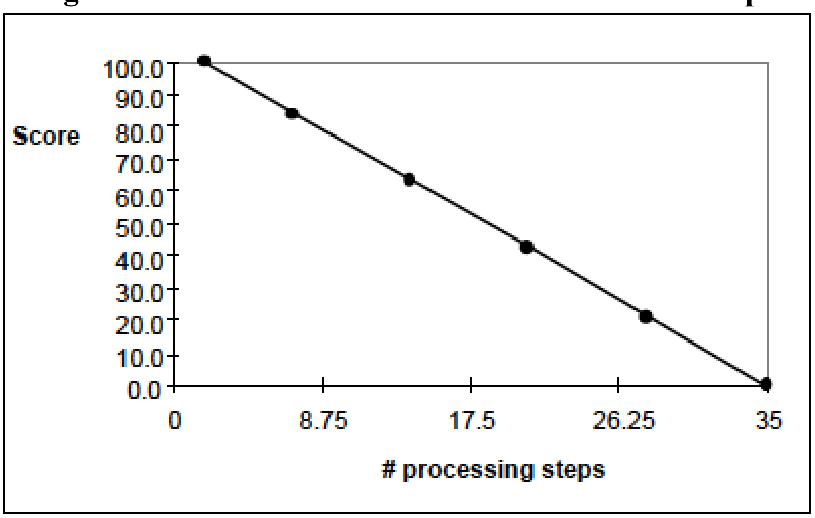


Figure 6: Value function for Maximum Plutonium Inventory

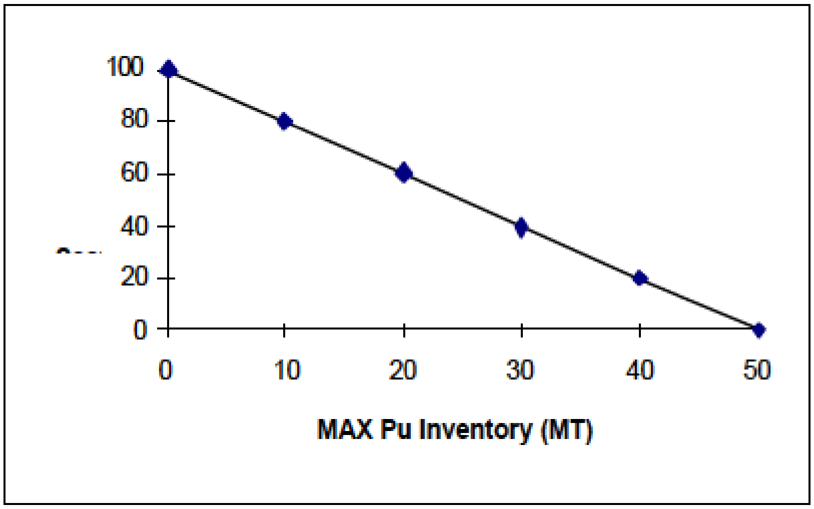
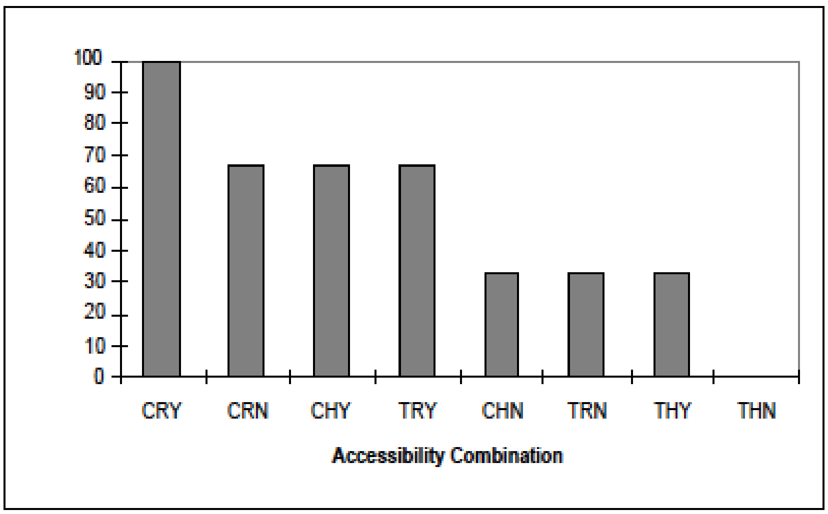


Figure 7: Value function for Accessibility of Material



The measures for Accessibility of material are based on the descriptions in **Table 1.** For example, the score CRY means that there is a tamper indicating container (C), the container must be accessed remotely or through robotics (R), and special equipment must be used to access the container (Y)

5.3.3 A functional form to combine scores to compute probability of adversary failure

For any step that the adversary must accomplish, there will usually be several features of the process and safeguards in place that bear on the likelihood that he will fail in executing the step. The more stringent the features the more likely that he will fail. A particular set of features may act synergistically to so that improving one feature enhances the effect of other features. They might also be relatively independent. It can work out that if one feature is sufficiently stringent, the probability of adversary failure is very high, regardless of the other features. These variations are analogous to the issues in combining features in a multi-attribute utility function (see Keeney and Raiffa, for example). We can use a functional form related to a multi-attribute utility function to capture these effects. Note that the proper calibration discussed in the next sections is essential to the validity of this interpretation.

To develop the basic equation, first consider the basic equation for the probability that an adversary will succeed when there are several distinct obstacles to overcome. This is [1-probability of succeeding at all obstacles]. If there are S obstacles, the probability that the adversary will fail is:

$$Prob\ Adv\ Fail\ = \left\{1 - \prod_{i=1}^{S} \left[1 - P_i^{fail}\right]\right\} \tag{1}$$

where:

P_i^{fail} = the probability that the adversary will fail at obstacle i
S = the number of obstacles

In our case we do not have distinct obstacles to overcome. There several features, either characteristics of the process, or safeguards that have been put into place that all affect the probability of failure. It is generally not the case that the adversary must actually "defeat" these features. Instead, it is more correct to consider these features to be "conditioning" events for the probability of the event "adversary failure". That is, the probability of adversary failure is a function of the values of these conditions. In these examples, it is clear that if all of the features are at their maximum level, the probability of adversary failure is increased, and if they are all at a low level, the probability of adversary failure in decreased. In addition, these features might, in some cases, work together, while in other cases they may act independently.

Without a detailed probability analysis, we expect to rely on the judgments of SMEs. It is expected that SMEs can mentally integrate many of these questions and provide a judgment about the effectiveness of the process features, and can provide an estimate of the effectiveness of combinations of features.

To use an SME, a model structure is needed that can capture judgments about the effectiveness of the individual features and combinations of features. It is observed here that the questions about the effectiveness of combinations of features are analogous to the questions about the interactions of attributes in a mutli-attribute utility (MAU) model. It is proposed here to use a probability model having a similar functional form. This model can be calibrated based on well-defined questions posed to an SME.

As an example, we construct the function modeling the probability that an adversary will fail to acquire the material dues to the difficulty of the task. Let the scores that describe the features be denoted x_{BT} (bulk throughput), x_{PS} (processing steps), x_{MP} (maximum inventory), and x_{AM} (accessibility of material). The value functions over the scores determined in **Figure 4** through **Figure 7** are denoted as $v_{BT}(x_{BT})$, $v_{PS}(x_{PS})$, $v_{MP}(x_{MP})$, $v_{AM}(x_{AM})$.

Using these functions we can define an overall function for the probability that the adversary will fail due to the difficulty of acquiring the material:

$$Prob\ Adv\ Fail\ Diff\ Acq = \frac{1}{K_{AcqDiff}}\left\{\prod_i \left[1 + K_{AcqDiff}\ w_i \bullet v_i(x_i)\right] - 1\right\} \qquad (2)$$

where

i = *BT, PS, MP, AM*

The components of this function are interpreted as follows:

- x_i is the score for the process or facility on feature i (e.g. Max Pu Inventory in MT)
- $v_i(x_i)$ is the value as a function of the score. Recall that this value ranges from 0 to 1 (or 0 to 100%). It is calibrated to be linearly related the probability of adversary failure as the value of x_i varies.
- w_i is a weighting factor for feature i . This can be interpreted as the probability that the adversary will fail when the value function for feature i is 1.0 (or 100%). This will be clearer when we discuss the procedure for calibrating the function.
- $K_{AcqDiff}$ is a calibration parameter which determines the degree to which the scores on the features interact. It is analogous to the K in a multi attribute utility function. It can greater than or less than 0. If it is large then no single feature can be very effective in defeating the adversary—they must all be working together to be effective. When it is negative, it implies that any one of the features by itself can play a strong role in defeating the adversary even if the other features are at very low levels. This will be illustrated in examples below.

5.3.4 Calibrate the functional form through expert elicitation

The parameters of the model are elicited through a set of questions posed to the SME. The elicitation procedure will first assess the values of the w_is and then assess the value of $K_{AcqDiff}$.

To assess each w_i we pose the following question to the subject matter expert (SME):

> "Assume that one of the features, say BT, is at its highest level so that $v_{BT}(x_{BT})$ = 1, and all the others are equal to 0. What is the probability that the adversary would be defeated in the phase of the stage (e.g. difficulty of acquiring material). "

The SME's assessment of the probability is the value of w_{BT}.

To illustrate, let us assume that we are assessing w_{BT}. If the value for BT is at its highest and the others are at there 0 level, the value of the function is

$$\frac{1}{K_{AcqDiff}}\{[1 + K_{AcqDiff}w_{BT} \bullet 1] \bullet [1 + K_{AcqDiff}w_{PS} \bullet 0] \bullet [1 + K_{AcqDiff}w_{MP} \bullet 0] \bullet [1 + K_{AcqDiff}w_{AM} \bullet 0] - 1\} \qquad (3)$$

which reduces to

$$\frac{1}{K_{AcqDiff}}\{[1 + K_{AcqDiff} w_{BT} \bullet 1] \bullet [1] \bullet [1] \bullet [1] - 1\} \tag{4}$$

or simply: w_{BT}

To complete the calibration, we must determine the value of $K_{AcqDiff}$. Its value is elicited by asking the SME the following question:

"Assume that all of the features are at their highest levels (all the *v(x)* equal 1.0). What is the probability that the adversary will be defeated?

Let the P_{total}^{fail} = the SME's estimate of the probability of adversary failure under this assumption. Under this assumption, the value of the function is:

$$P_{total}^{fail} = \frac{1}{K_{AcqDiff}}\left\{\prod_i [1 + K_{AcqDiff} \bullet w_i \bullet 1] - 1\right\} \tag{5}$$

or, expanding out

$$P_{total}^{fail} = \frac{1}{K_{AcqDiff}}\left\{\begin{matrix} [1 + K_{AcqDiff} w_{BT}] \bullet [1 + K_{AcqDiff} w_{PS}] \\ \bullet [1 + K_{AcqDiff} w_{MP}] \bullet [1 + K_{AcqDiff} w_{AM}] - 1 \end{matrix}\right\} \tag{6}$$

Analogous to a multi-attribute utility function, we calibrate the function by solving for the value of $K_{AcqDiff}$ that satisfies this equation.

In the case of a MAU function, the K is calibrated such that the overall function is 1.0 when all of the attributes are at their highest level. This is because utility functions are defined to range between 0 and 1. An overall utility of 1.0 implies that all attributes are as good as they can be. In the case of modeling probabilities, we note that even when all features are as good as they can be, there is still some probability that the adversary will be successful, so the probability of adversary failure is less than 1.0. In the probability case the value of $K_{AcqDiff}$ is calibrated so that the function equals P_{total}^{fail} when all of the features are at their highest value.

5.3.5 Calibration provides insight into interdependencies of attributes

The probability model rolls up the effects of the various features of the process to model the probability of adversary failure. However, making the assessment implies assumptions about the interaction of the features of the process that determine the overall probability. The magnitude of K in relation to the sum of the w_i s reflects the assumption made by the SME regarding the interaction of the features.

The sum of the w_is is significant, just as in MAU. But, in the case of an MAU, the sum of the w_is is compared to 1.0, which is the maximum value that the utility function can have. If the sum of the w_is is greater than 1.0, the attributes are *substitutes*. If the sum is less than 1.0, the attributes are *complements*. If the sum is equal to 1.0, the attributes are *additive*. When attributes are substitutes, the maximum utility can be reached even if all of the attributes are not at their maximum value. Put another way, if one attribute is high, improving another attribute does not improve the overall utility as much as in the additive case. Conversely, when the attributes are complements, all of the attributes need to be high in order to maximize the overall utility. Putting it the other way, if one attribute is high, the effect of improving one of the other attributes is magnified compared to the additive case.

In the probability model proposed here, there is an analogous effect. Here the test is the sum of the weights compared to P_{total}^{fail}. There are four cases of interest:

- The sum of the w_is is less than P_{total}^{fail} then $K_{AcqDiff}$ >1
  This implies that the features must work together to defeat the adversary. Simply having one feature working well is not adequate to maximize the probability of adversary failure. Put another way, when one feature is strong, the effects of the other features are increased.

- The sum is equal to P_{total}^{fail} then $K_{AcqDiff}$ →1
  The features do not interact with each other, either positively or negatively.

- The sum of the w_is is greater than P_{total}^{fail} then $K_{AcqDiff}$ <1
  This implies that the features are more or less independent. When one feature is strong, the effect of improving the features is reduced.

- The sum of the w_is is greater than P_{total}^{fail} and P_{total}^{fail} is less than, the largest w_i, then $K_{AcqDiff}$ <1
  This case seems to be problematic. It implies that one of the features alone (the one with the largest w_i) would be preferable to having all of the features working together. If the feature with the largest w_i is at its maximum and the others are at their minimum, the probability of adversary failure is w_i. However, if the other features are brought to their maximum strength, the probability of adversary failure *decreases* to P_{total}^{fail} . This case does not seem logical. If an elicitation produces this result, the logic of the situation should be reviewed.

5.3.6 Summary of process

The full analysis is executed by developing models of adversary failure at the lowest level using the functional form defined in equation (2). Then computing the probability of adversary failure for each stage, and the probability of failure for all of the stages. The analysis is completed in the following steps:

1. Assess the probability of adversary failure due to a) difficulty and b) detection for each stage
2. For each stage, compute the probability of adversary failure by combining the probabilities of adversary failure due to difficulty or detection using the probability form as in equation (1)

3. Compute the overall probability of adversary failure by combining the probabilities of adversary failure at each of the stages using equation (1)

The structure of this model assumes that the relationship between the probabilities is linear. The probability contribution from a single feature is $w_i \bullet v_i(x_i)$. Recall that $v_i(x_i)$ is scaled from 0 to 1 and w_i is the probability of adversary failure when the feature is at its best state and the other features are at their worst state. Thus, $w_i \bullet v_i(x_i)$ is the probability contribution of feature i, as a function of the state, x_i.

The overall probability is always a linear function of $v_i(x_i)$ as x_i is varied. As discussed above improvements on one feature can affect the impact of improvements on other features. For example, in the case of complementary features, the usefulness of feature *i* improves if feature *j* is present. Therefore, if the state of feature j is high, the slope between the overall probability and the value of feature *i* will increase—improvements in feature *i* will have more value if feature *j* is at a high state.

When the states of all the features are at the 0 level, the probability function will produce a value of 0. This indicates that the probability of adversary failure is 0. The physical situation corresponds to a case where the facility is un-locked, un-manned, and un-monitored. This is not usually the case for any facility. However, it is important the value functions be defined such that their lowest level does describe a facility such that the probability of adversary failure is 0. This will provide the proper baseline values for comparison of proposed features. Clemens and Smith discuss the importance of correctly specifying the baseline when portfolios of upgrades are considered.

The method proposed here is developed to estimate the probability of adversary failure in and attempt to acquire material and make a device from it. At the higher levels of the analysis, the structure is well enough defined to use standard probability models. However, at the lowest level of the analysis standard probability models cannot easily be developed. This paper suggests a method to use elicitations from SMEs to calibrate a probability model that incorporates the effects of the different features of the process and material. Although is may be expedient to use this approach in many situation, the probability assessments could be derived from rigorous methods. For example ASSESS [Al-Ayat 1993] develops a model of the processes, facility, and safeguards. It uses a database of probabilities of detection and the success of adversary actions. It can use these to compute the probability of success for any specified adversary strategies. Rigorous probabilities of adversary failure could be computed using models such as these.

The proposed method has been formulated to address the concerns about the structure and application of evaluation methods for proliferation resistance.

- This method addresses a specific, concrete element of the proliferation resistance analysis: the probability that and adversary would fail to acquire material and process the material into a device. This is a fundamental building block of analysis of proliferation resistance or proliferation risk

- The method is based on a well-defined metric: probability of adversary failure. Given this definition, the metric can be used to objectively compare different fuel cycles or different configurations of a fuel cycle. This definition makes it possible to evaluate the cost benefit trade-offs for assessing alternative cycles or configurations
- The method is practical to use. The most difficult part of the analysis is modeling probabilities of adversary failure at the lowest level of the model where the effects of different features of the cycle of the facility are not well defined. The model allows for a reasonable subjective probability assessment from SMEs. The calibrating questions are well defined. The model combines these in a reasonable way. The probability assessments could alternatively be done through rigorous modeling. This is a reflection of the fact that the assessments are well defined.
- The method is transparent and diagnostic. An analyst can identify the features of the process that contribute the most, or the least to the probability of adversary failure. The subjective probability model is transparent in that an analyst can determine the assumptions regarding the interaction between the features of the process that contribute to the overall assessment.

6.0 Summary and Conclusions

We believe that the NRC study did not select the best examples of multiattribute utility analysis (MAU) studies for their evaluation of the process. Many of the shortcomings in execution of the studies reviewed could have been overcome by using rigorous, accepted methods commonly used by the decision analysis community. As the NRC noted, the studies reviewed were not conducted by experienced decision analysts.

A multiplicative MAU form might be more effective for this application due to its similarity to a probabilistic model of threats to the nuclear fuel cycle. One implementation of such a model is developed in this white paper.

References

[Al-Ayat 1993] Al- Ayat, R. A., Analytic System and Software for Evaluating Safeguards and Security – ASSESS User's Manual, UCRL-TR-223631 (1993)

[Butler 2005] Butler, John, Alexander Chebeskov, James Dyer, Thomas Edmunds, Vladimir Oussanov and Jianmin Jia, The Adoption of Multi-attribute Utility Theory for the Evaluation of Plutonium Disposition Options in the United States and Russia, *Interfaces Vol. 35, No. 1*.

[Clemens 2009] Clemens, R. T., and J. E. Smith, "On the choice of baselines in multi-attribute portfolio analysis: a cautionary note", *Decision Analysis*, *Vol. 6, No. 4*, pp 256-262, December 2009.

[Dyer 1997] Dyer, James S., Thomas Edmunds, John C. Butler, and Jianmin Jia, Evaluation of Alternatives for the Disposition of Surplus Weapons-usable Plutonium, ANRCP-1997-1, Amarillo National Resource Center for Plutonium (1997).

[Dyer 1998] Dyer, James S., Thomas Edmunds, John C. Butler, and Jianmin Jia, A Multiattribute Utility Analysis of Alternatives for the Disposition of Surplus Weapons-grade Plutonium, *Journal of Operations Research, Vol. 46, No. 6*.

[Edwards 2007] Edwards. Ward, Ralph F. Miles, Jr., and Detlof von Winderfeldt, *Advances in Decision Analysis: From Foundations to Applications*, Cambridge University Press, New York.

[Garcia 2012] Garcia, Humberto E., et al., Integration of Facility Modeling Capabilities for Nuclear Nonproliferation Analysis, *Progress in Nuclear Energy, Vol. 54*, pp. 96-111.

[Keeney 1976] Keeney, Ralph L. and Howard Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley & Sons, New York.

[Keeney 1992] Keeney, Ralph L., *Value-Focused Thinking: A Path to Creative Decisionmaking*, Harvard University Press, Cambridge.

[Maurer 2009] Maurer, Stephen M. (editor), WMD Terrorism: *Science and Policy Choices*, MIT Press, Cambridge.

[Ni 2013] Ni, Kevin, Daniel Faissol, Thomas Edmunds, and Richard Wheeler, Exploitation of Ambiguous Cues to Infer Terrorist Activity, *Journal of Decision Analysis, Vol. 10, No. 1*, (March 2013).

[NRC 2013] National Research Council, *Improving the Assessment of the Proliferation Risk of Nuclear Fuel Cycles*, National Academies Press.

[Ward 2007] Edwards, Ward, Ralph F. Miles, Jr., and Detlof von Winterfeldt, Eds., Advances in Decision Analysis, Cambridge University Press (2007).

[Winterfeldt 1993] von WInderfeldt, Detlof, and Ward Edwards, *Decision Analysis and Behavioral Research*, Cambridge University Press, New York.

Appendix: Selected comments from National Academy of Sciences report

This Appendix reproduces sections from the 2013 National Research Council study titled *Improving the Assessment of the Proliferation Risk of Nuclear Fuel Cycles.* These sections are particularly relevant to this white paper.

1. (ES-2) The committee has identified the following applications as opportunities in which the current predefined frameworks can provide value and utility to policy makers if the shortcomings in their execution are addressed:
   - Comparing the proliferation resistance of fuel cycles and identifying locations to apply safeguards or material monitoring,
   - Educational applications (e.g. academic applications or new nuclear energy states), and
   - Enabling consistent communication with international partners or the public on nuclear energy decisions by providing analysis through the application of a predefined, internationally accepted and known methodology
2. (ES-2) Nonetheless, the committee does not support a new or expanding R&D program in this area. Based on discussions with policymakers, the committee determined that the existing tools have limited utility to inform their nonproliferation decisions beyond what a case-by-case analysis would produce. Case-by-case analysis also uses expert knowledge and can suffer from the same challenges listed above regarding predefined frameworks. However, their use of expert opinion and knowledge is clear and understood by policymakers while predefined frameworks' use of expert knowledge is less clear because it is often combined and presented as an integrated result.
3. (ES-2) The committee recommends that fuel cycle R&D decisions include proliferation resistance instead of proliferation risk as one factor among others (such as cost and safety) to guide those decisions. Technical assessments are limited by the availability of technical details associated with future nuclear fuel cycles. Therefore, the committee recommends that DOE-NE and NNSA jointly decide upon a set of high-level questions comparing the proliferation resistance of proposed future fuel cycles to the current once through fuel cycles. Assessments should be revisited at key milestones throughout the technologies' development and eventual deployment as new and better information and data emerge.
4. There is a wider range of issues faced by policy makers, e.g. the probability that a host nation would attempt to proliferate
5. (pg. 9) However, the committee found several examples in other domains within the U.S. government in which decision makers use predefined framework-like tools to inform decisions. Examples include prioritizing countries for engagement on nuclear, chemical, and biological security (the Office of Cooperative Threat Reduction within the Department of State; Dolliff 2012), and using risk assessment methods for optimizing architectures for global nuclear detection (the Domestic Nuclear Detection Office within the Department of Homeland Security; Streetman 2012). These examples show that some policy and decision makers do find utility in frameworks that can deconstruct complex problems into their component parts. In these instances, the policy and decision makers were actively involved in the analysis process and not interested only in the final results. The frameworks provide a

structure for organizing complex problems with a large number of variables and assessing which factors are most important to the results.

6. (pg. 10) The committee considered a set of predefined frameworks and found the following shortcomings in their execution:
   - poor and/or undocumented expert elicitation processes, and
   - lack of sensitivity and uncertainty analyses
7. Inherent limitations of applicability included:
   - for future fuel cycles, unknown facility and host-state details, and
   - limited shelf life of assessments
8. (pg. 11) RECOMMENDATION 3.1: DOE-NE and NNSA should consider whether elements of a formal PRA approach could improve multidisciplinary assessments of proliferation risk, especially the quantification of uncertainties. Although the committee concluded that work on understanding motivations to develop nuclear weapons and modeling an adaptive adversary do not have evidence-based records of success in real-world situations, it supports the inclusion of such approaches into proliferation risk analysis when and if they have an established quantitative basis.
9. (pg. 24) The committee notes that assessment of proliferation risk of a particular nuclear fuel cycle includes analysis of country-specific factors such as the probability that a country's leadership would choose to proliferate using a specific fuel cycle and the probability that it would be successful if it chose to do so. The proliferation resistance of the fuel cycle in question could contribute to both of these factors. Potential questions about proliferation resistance of potential future fuel cycles were discussed and determined by the committee to include
   - Are there significant differences in resistance to proliferation (e.g., time, cost, physical barriers, safeguard-ability, or transparency) associated with different potential future fuel cycles compared with those that exist today?
   - Can extrinsic measures such as physical security and international safeguards, intrinsic measures such as reactor design or material composition, or new operational concepts significantly increase resistance?